Inform Stor. Retr. Vol. 3, pp. 219-291. Pergamon Press 1967. Printed in Great Britain

THURSDAY, SESSION IV

SOME REMARKS ON INFORMATION LANGUAGES, THEIR ANALYSIS AND COMPARISON

DAGOBERT SOERGEL

CONTENTS

A. General remarks on information languages
A.0. Use of the term "Information Language" (IL)
A.1. Purposes of an information language—requirements for an information language
A.2. A general model for the structure of the ensemble of subjects
A.3. Specializations of the general model
4.1. Indicative and informative descriptions
4.2. Analytic vs. synthetic relations
4.3. Paradigmatic vs. syntagmatic relations
A.5. Synthetic relations, roles, and categories
B. Comparative list of some schemes given in the literature
B.1. Principles of construction and state of work
B.2. Some feature evolving from the comparison of some schemes of role indicators
B.3. Cumulative and comparative list
B.4. Sources of relations, role indicators, and categories
References

A. GENERAL REMARKS ON INFORMATION LANGUAGES

A.O. *Use of the term "Information Language"^)*

We take the following terms as synonymous:

Information language, information retrieval language, information description language, documentary language. Sometimes it is also called machine language (Upenskii, 59:357). There are two possible confusions which occur frequently and which one should be cautious to avoid:

0.1. "The machine language [information language in our terminology D.S.] should not be confused with the machine code" (Upenskii, 59:357). In Ranganathan's more general terms: There should be no confusion between the idea plane and the notational plane. Many difficulties in classification arise from this confusion. This paper is devoted mainly to problems on the idea plane.

0.2. "The IR-language I want to talk about, is the language in which the information to be retrieved is formulated, i.e. a declarative sentence language, not the imperative sentence language of the programmer-retriever" (Bohnert in Sammet, 62.1:10), that is, the information language should not be confused with the programming language. (To avoid misunderstandings it would perhaps be better to say "descriptive" instead of "declarative" and to state explicitly that in a descriptive sentence an imperative may be described which is> however, not to be executed by the system but to be retrieved by appropriate searches.)

Finally we shall exclude terminological problems from the main concern of this paper, that is, we shall assume that the information to be described by the IL is completely known

or, to put it another way, that the difficulties arising in a natural language by synonyms and homonyms have already been resolved.

A.I. *Purposes of an information language—requirements for an information language*

(In the following discussion we rely mainly on Gordonnier, 59/61:1093-1102, Kochen, 65, Melton, 58a, Ohlmann in Sammet, 62.1:10, Upenskii, 59:357-59. These sources are cited by the names only; if needed the exact page is given.)

In short we may say that the purpose of an IL is the description of information for storage, retrieval, processing, and communication. Each of these functions should be formalized as far as possible because ".. .everything which can be formalized can be made automatic" (Upenskii, 358, also point 2), on p. 359). Let us look at these points in some more detail:

I. The *description* should be possible at various levels of specificity, and it should be possible to produce a less specific one automatically from a more specific description (Ohlmann's point "(1) condense material (digest, abstract, and index)"). It need not be mentioned that the levels of specificity vary from the description of single facts or other information units to the assignment of markers (subject headings, index terms) to some groups of facts or other information units. The description of information is made in order to perform the following functions. II. *Processing of information* or perhaps in more appropriate formulation "(3) relate or associate material (place isolate facts in context)" (Ohlmann). One important mode of such processing would be "(2c) deductive reasoning" (Upenskii, 359; the same point is made by Kochen, 65, by Bohnert, 63.12, and by Levien, 65.12); one may differentiate strict logical inferences which are possible in a calculus of logic (compare Bohnert, 63.12) and plausible inferences (Levien, 65.12:77). The processing may be done before storage, during retrieval or after retrieval.

III. *Storage*

The store must be a dynamic one so that it is possible to "(2) modify material (add, delete, and reorder)" (Ohlmann). It may seem that this point has mainly programming aspects, but this is not so:

(a) Before a new information unit is added one must first see if this information is already contained in the store, in other words, the "(b) identification, in various ways, of recorded facts [more general: informations D.S.]" (Upenskii, 359) is needed (this problem is related to retrieval, see IV.).

(b) In deleting information one must give due consideration to the possibly interconnected structure of the store (for example, due to IL).

IV. *Retrieval*

It goes without saying for all authors that the IL must be adapted to appropriate retrieval, including retrieval for selective dissemination. It has been widely discussed that this implies (a) the introduction of relevance-constants and (b) the possibility of varying the scope of the search question (performing inclusive searches). It is also well known that the information which is asked for in spite of being not explicitly stored may be deduced from the information in the store (see IL).[(1)]

[(1)]This point is equivalent to Ohlmann: "(4) translate (match, requests, machine capabilities, and .answers)", which includes also point (7) below, and to Upenskii: "(2(a)) the selection of necessary information" and'partly to Melton's (3) (see below, V.(7)).

V. *Communication*

Communication comprises both communication between man and man and between man and machine (as needed for information retrieval for instance). Cordonnier has made a proposal to create a common language U for communication. This language U would comprise all the possibilities for the differentiation of expressions contained in the source languages but would eliminate Synonymity and Homonymity. A poorer version of this language called M would be abstracted for use in machines. Freeman, *65,* also proposes to write scientific papers in the social sciences in an artificial information language because natural language achieves only partial and not full communication; this results from the fact that different people assign different meanings to the same terms without making those differences explicit.

Iii this connection one may mention the idea of using an IL as a metalanguage for machine translation (Cordonnier; Upenskii, 359/60; Melton, 58e:337/38). The claim of UDC that for each concept as expressed by the UDC number one may find the corresponding terms in different languages by consulting the respective editions of UDC lies along this path, too.

Now, in order to fulfill these functions in a reasonable way an IL should have the following features:

(1) It "(1) must be *unambiguous.* Every description should permit of an un equivocal interpretation". (Upenskii, 359.)

(2) "It should *be flexible.*[79] (Kochen.)

(3) It should be "£«#/M[Freneh; ... D.SJ-like" (Kochen) or be easily trans formable to an English-like version.

(4) It should be adapted to "(1) the nature of the subject matter..." (Melton). It may be that "specific abstract languages.. .[for].. .the various fields of sciences" (Upenskii, 359) have to be created.

(5) It should be possible to detect some kinds of incompleteness of information (see Bohnert, 63.12:52) and incorrectness of information, 'correct' being interpreted in a purely formal manner as 'following the specified rules of the information language which may reflect some ways of thinking which by common agreement are thought necessary for a given field of science' (Cordonnier). (Compare A. 5.4.)

These points (which perhaps together make up Melton's point "(4) the semantic capabilities of the system of encoding") must be brought into accordance with the following additional "engineering" requirements:

(6) "(2) the nature, scope and variety of information requirements." (Melton.)

(7) "(3) the special requirements of automatic searching [perhaps more appropriate: processing D.S.] by electronic equipment" (Melton) or other equipment, as may be the case.

(8) Finally "(5) the time, cost and effort involved in the total operation" (Melton) is an important factor.

It may be that it is not possible to fulfill all these requirements with one language in one version. It would then be necessary to have different versions of the same language, each version adapted to one or more purposes, so that each version may be easily transformed into any other. A general proposal in this direction was made by Qordonnier (see above); Bohnert, 63.12, describes a project, the purpose of which is the creation of some restricted form of natural English which may be transformed into the language of a calculus of logic.

A.2. *A general model for the structure of the ensemble of subjects*

(Regarding those aspects of the subjects which are of importance for the purposes mentioned above.)

This is a somewhat modified summary of a more detailed analysis contained in Soergel, 65.9.

2.1. *Generation of the ensemble of subjects (expressions)*

We start from a set of ($A^{+\,+}$) of primitive concepts (predicates, attributes) and a set M of "relevance constants". As usual more complex subjects are described by combinations of simpler subjects. This combination is achieved as follows: first to each of the more simple subjects a relevance constant is assigned in order to specify the importance of the more simple subject in the frame of the more complex subject to be described or, to turn the argument around, to specify the relevance of the complex subject for searches for material on the more simple subject. Then a binary operation "CONNECTION" is defined which allows for the construction of a complex subject from two more simple subjects to which relevance constants have been assigned. This operation requires the indication of a further constant, the "Connection strength" which for the sake of simplicity may be an element of the set M. We generalize the definition of this operation in allowing that to one subject there may be connected different other subjects at the same time, each in a binary way. In short the operation CONNECTION allows for the construction of a network representing a complex subject, the points of the network being more simple subjects and the lines between the points being weighted. Of course the points contained in the network may be networks in themselves. To give a simple and well known example:

Thermometer would be represented by "instrument :{purpose :[(operation measure):(< target measure > -.temperature)]}".

By recursion one arrives at the primitive concepts; the basic set (A^{++}) generates the whole ensemble of possible subjects. We call each subject derived in such a way an expression of (= within) the information language.[(1)]

2.2. *Relations between subjects*

Before analyzing these expressions of the language in more detail we shall finish the description of our general model by looking on the relations *between* subjects. The statements on these relations may be called meta-statements about the information language. We may roughly differentiate three types of such relations between subjects A and B:

(1) A partial ordering relation A <= B for which the best wording is perhaps "A implies B". This relation may hold (a) between two subjects at least one of which is a primitive concept due to stipulation, the reason for which lies outside of the system of the IL (for example biological taxonomy: horse implies mammal); or (b) between two subjects at least one of which is not a primitive concept due to some rules of inference which have been defined for the information language. Such a rule would be for example A <= B if B occurs as a point in the network of A, for example: A = Thermometer, B = temperature or B = (operation: measure).

The whole ensemble of possible subjects becomes such a partially ordered system.

(1) In an exact formulation, we should distinguish symbols and concepts represented by the symbols; an expression would then be a string of symbols, created according to specified rules, and would represent a subject.

(2) The second type of relations would be the nearness of two subjects A and B in the partially ordered system described in $(1)_5$ which may be measured in different ways (compare Soergel, 65.9, chapter 4).

(3) The third type of relations may be characterized as "see-also-relations". They hold either (a) between any two subjects by stipulation or (b) between two subjects at least one of which is not a primitive concept by rules of inference. Such a rule would for example be if "A see also B" and if one gets the network of B' from the network of A' by replacing A by B, then "A' see also B"\ For example: "Instruments for the measurement of tempera ture" see also "Instruments for the control of temperature".

2.3. *Relational factors; role indicators*

Having finished the general frame of our model we may proceed to a more detailed analysis of the expressions of the IL which will bring us to the relational factors in information retrieval From a pragmatic point of view it might be useful to divide the primitive concepts into "content concepts" and in "role indieators"[(1)] (of course this division will be questionable in single cases). The smallest building bloc of a network would then be a subject resulting from a connection of a content concept with a role indicator (such a connection is sometimes called role-term). The connection of two such role-terms would then be equivalent to a binary relation applied on the two content concepts. *Higher order relations* are built up in an *obvious* way.[(2)]

If one has a set of specified higher order relations and one wishes to express these relations by the IL one has to go the other way around and to break down the higher order relations into role-indicators which then may be incorporated into the IL.

Of course role-indicators may be connected with other role-indicators to give more complex role-indicators. Between two role-indicators all the relations specified above may hold. The search for a basic set of role-indicators to which this symposium may contribute differs not, in principle, from the search for a basic set of content concepts. The rigorous application of the principle "breaking down to role-indicators" will be useful in this connection as is to be seen in part B.

2.4. *Remark on generality*

I realize that there are not many, if any, new ideas in the model described above but I hope that it has some significance by the very fact that it lacks some features of other systems also called general models. Such additional features are due mostly to machine requirements or other artificial and accidental reasons; for example; distinct levels of complexity of subjects (numbered from 1 to 4, etc.), distinction between subjects contained in the vocabulary and those that are not contained there, distinction between analytic and synthetic relations.

2.5. *Engineering aspects*

Now the abstract model described gives rise to a structure of high complexity which cannot be easily handled either by man or by machine. In order to get information languages for concrete applications one must therefore introduce some restrictions. In doing so one should be eagerly aware of the information losses which result as a consequence. Section A.3. is concerned with this problem.

(1) Definition of these classes may be by enumeration.

(2) As may be seen from the last sentence it is not stated here that relations are ever binary.

2.6. *Concluding remark*

The general model described in this section may be interpreted in two ways:

(a) if one goes back to the "really" primitive concepts as the elements of the set (A^{++}) one has a model for the "real" structure of subjects.
(b) It is possible to model any existing system for the expression of subjects—including probably natural language—by taking as elements of (A^{++}) those entities (terms, concepts, or whatsoever) which are defined as primitive in the scheme to be modeled (compare Soergel, 65.9, chapters 1 and 2). Comparison with (a) shows the short comings of each of the systems.

A. 3. *Specializations of the general model*

There are two main possibilities for simplifying the information language:

(a) *Drop the "links"* that is, the information that two subjects are connected. For searching this means that one has to state as connected any two subjects contained in the information unit (for example, in a document description).
(b) *Drop the role-indicators.*

These and other possibilities will be discussed in the following:

3.1. *Coordinate indexing*

Restrictions (a) and (b) are introduced. It is possible to *reintroduce links* to some extent in a system of coordinate indexing in two ways:

(i) "In systems without links, links may be simulated to a certain degree by changing the set of objects (documents; information units). More properly speaking if the attributes to be linked belong to a certain section of a document one may separate this section and define it as an own document in itself. Thus the attributes applicable to this section are not connected with the attributes of other sections. In this way it is even possible to simulate different discrete linking strengths by different.. .[subdivisions of information units]. But one should mention that this possibility is very restricted, because it often occurs that the terms to be linked do not belong to a separate section" (Soergel, 65.9:1.125, p. 17).

(ii) One may stipulate some rules stating that subject A can be connected only with such and such other subjects. This procedure, which I term "indirect prelinking", is thus possible in systems without role-indicators, and is of special importance in systems with role-indicators where one may formulate the rules for the indirect linking of role-indicators with certain classes of terms, or of role-terms with role-terms, considering only the roles contained.

Because the role-indicators are few in number and general in application, a few rules will do the job for many cases (see A.3.3.),

3.2. *Systems with links, without role-indicators*

Some procedures of statistical indexing deliver for each document a set of index terms together with some statistical connections between these terms. Information languages for human indexing of this type would also be conceivable. It has been stated that links convey more information than role-indicators (see A.3.5.).

3.3. *Systems with role-indicators, without links*

(i) Such systems seem not to be useful if they do not introduce some device of indirect prelinking (see A.3.L (ii)).

Example: in a German documentation on problems of developing countries there exists the agent/target problem in the description of international relations; in most cases there is involved one developing country and one industrial country. So we introduce the "pre-linked" role-indicators "agent developing country", "agent industrial country", "target developing country" and "target industrial country". Of course this device does not work if two countries of the same type are involved, the one being agent, the other being target. We may interpret this as follows: To the role-indicator a specification is added, specifying to what kinds of elements the role-indicator may be connected in order to get a role-term.

(ii) One may weaken the restriction "without links" and allow for direct linking of role-indicators with content concepts, but not for further direct linking of one role-term with another. In this case, an indirect pre-linking is automatically achieved by correct splitting of synthetic relations and taking all the role-indicators. For example (compare B.3.); relation: Property/Having the property; takes both "Property" and *"Having the property"* as role-indicators. In establishing the connection within an information unit one may then drop those which do not result in appropriate relations, that is, role-terms containing "Property" may only be linked with role-terms containing "Having the property".[(1)] This device is realized by the WRU-role-indicators:

"KWV property" and "KOV property given for" (more appropriate: "Having the property"; KOV is restricted in the WRU-System to materials, but this is an unnecessary restriction). To give a further example: "KUG" (location given for) is used with the name of a company or other organization or group of individuals *to ensure proper designation of location* when the role-indicator KIG (location) is used" (Rees, 58:186). The connection organization/location is established in an unique manner as long as more than one organization and more than one location do not simultaneously appear in the same information unit. A further type of indirect prelinking of role-terms may be introduced as follows (compare BA Ky, 61):

Divide role-indicator A into A' and A" and stipulate that role-terms using A' may be connected only with role-terms of such and such categories (see A.4.3.), and accordingly for A".

3.4. *The Semantic Code^ and telegraphic abstracts of WRU*

This information language exposes various features already described. It has an artificial division of subjects in levels: the semantic factors (primitive concepts), terms, phrases, sentences, paragraphs, etc. These levels are used for the simulation of links. Role-indicators may be connected by direct linking with primitive concepts (semantic factors); the combination of several such factors, each linked with a role-indicator, make up a term. A term in turn may be directly linked with a role indicator, and a combination of such terms makes up a phrase. Phrases and higher units of information may not be combined with role-indicators and serve only the purpose of simulation of links. There are two kinds of role-indicators: the ones used in connection with primitive concepts and called analytic relations and the

[(1)] In systems with links, "Having the property" is obviously not needed as role-indicator. [(2)] "Semantic Code" mixes two things which according to A.0. should be strictly separated, namely the semantic information language and the machine code, by which this IL is represented in the machine.

others used in connection with terms and called synthetic relation (see A.4.2.). Some of the role-indicators have an indirect prelinking function (see the examples in 3.3.).

3.5. *Discussion of the practical importance of role-indicators in information retrieval* Doyle, 63.2:11, has made the point that if some terms (or in our terminology content concepts) are associated (connected), then it is very likely that the roles of these concepts correspond to a certain pattern. To cite a famous example; if the terms man, dog, bite are associated then it is likely that the subject is "dog bites man" (that is dog is agent, bite is action, man is target), and it is very unlikely that the subject is "man bites dog". Because this is true for the description of documents and of search questions it is very likely that the roles of associated terms are the same in both cases, so that in performing the search we need not know what the roles are. Doyle concludes "that one can create for himself a non-existing problem by confusing what *can* happen in a retrieval system with what *is likely* to happen". In commenting on this argument of Doyle we must distinguish between indicative and informative statements (see A.4.I.).

(i) Systems using indicative statements (only document retrieval systems).

(a) There are fields and combinations of terms for which Doyle's point does not hold. The problem is quite common in international economics (the combination Germany, United States, export); or in chemistry (reactants and reaction product).

(b) Doyle's argument may hold, but in spite of this there may be some infrequent cases in which the roles do not correspond to the standard pattern. There may be for example cases of "man bites dog". If someone studying psychology has the search question "man bites dog" he must go through all the frequent cases "dog bites man" in order to find answers.

(c) In various fields-especially in social sciences—there occur search questions which combine very general terms in specified roles. In this case irrelevant material would be too frequently found by dropping the role-indicators.

To cite an example from natural science (lacking concrete data on social sciences), in the Semantic Code there are 398 different concepts in which "material properties" is contained as a semantic factor, but there are only 28 different concepts in which this same semantic factor occurs as object of an action. (Melton, 58d:239). One might perhaps use in this case Doyle's argument to proceed as follows: assign to each combination of concepts a standard pattern of role-indicators and add these role-indicators each time this combination of concepts is found. However, this would have the disadvantage that the exceptional cases which correspond to the search question by virtue of exceptionality (and which therefore are perhaps the most interesting ones) are not found.

As a consequence we may say that the applicability of Doyle's argument is dependent on the nature of the field and the type of search questions. (d) Finally one may argue (as for example Barhydt, now at WRU, does) that it is better to have a list of specific terms, so that the description of information at a specific level is possible without role-indicators complicating the indexing process. In fact this means to go back to subject headings, the shortcomings of which for searching are well known (compare A.I., point IV.). The reason

for the shortcomings lie in the fact that the subject heading lists are lacking in proper designation of semantic relations between the subject terms. To resolve this lack, i.e. to organize the subject heading list, some approach realizing the general model described in A.2., which may include the use of role-indicators, is necessary (compare Soergel, 65.5).

(ii) For informative statements Doyle's argument is not applicable at least in principle because we cannot mix ordinary cases (which by the way are information-poor) with exceptional cases (which are the information-rich ones). One may introduce however some sort of appropriate coding of the ordinary cases by omitting the role-indicators and saying, if this combination of concepts occurs and if no role-indicators are specified, that the standard pattern of roles assigned to this combination of concepts applies.[1]

4.1. *Indicative and informative descriptions*

The statements "measurement of temperature" and "temperature was measured'* would have the same description in the IL outlined in our general model, and yet are different in nature. The first phrase describes a problem without saying whether there is a solution or a positive answer. The statement "It was not possible to measure temperature" or even "It is impossible in principle to measure temperature" (a wrong statement, of course) would fall in the range of the problem "measurement of temperature". The second statement also describes a problem (if implicitly), and gives furthermore a result, namely: temperature *was* measured. So it is not as much a problem of how the expressions of our IL are constructed as of how they are interpreted. We may add to our language some symbols to be prefixed to expressions in order to specify how these expressions are to be interpreted.. Such prefixes would be for example: *indicative*—states a problem,
informative—gives a result (and perhaps states implicitly a problem).
normative—states something to be done.

But these definitions would not completely correspond to the usual meaning of (for example) "indicative abstract", because "temperature has been measured" would be a sentence of an indicative abstract whereas an informative abstract has to give in addition the degree of temperature which has been measured. Now it is clearly to be seen that in the phrase "temperature has been measured" something is lacking which has to be given in each measurement, namely the value (or values) of the measured. Thus there is an empty place (variable would be misleading perhaps in this connection) in this subject which has to be filled in. So we may modify our definition as follows: an indicative expression is either an expression in the indicative mode or an expression in the informative mode with an empty place to be filled in. Such an expression in the informative mode stands for many possible expressions according to the various possibilities for the thing to be filled in. Regarding the partial ordering of the ensemble of all subjects we may say that the expression in the informative mode implies the same expression in the indicative mode, and that an expression in the informative mode with an empty place is implied by the same expression with an empty place filled in. But if we introduce the degree of specificity of a subject (for example the length of the chain from this subject to a maximal point of the partially ordered system),

[1] One should be cautious not to misunderstand this point. Roles are applied to the standard cases as well as to the deviant ones. Only the coding is different: For the standard cases coding is implicit by prior convention; for the deviant cases coding is by explicitly naming the roles.

there may be expressions in the indicative mode with a higher degree of speciality than other expressions in the informative mode: take for example the indicative phrase "measurement of temperature above 1000°C with constantan/copper/thermo-couples" and the informative phrase "the temperature has been measured to be 10°C".

It is clearly to be seen that there is no strong borderline between indicative and informative expressions because the set of "things" to be filled in in the empty place may be narrowed down step by step, making the indicative expression more and more informative.

In spite of the fact that specificity and informativeness of a subject or of an abstract must be distinguished, it is quite obvious that informative expressions are only useful if a certain degree of specificity may be reached in the IL. So the intention of being informative is one reason among others to have an IL capable of special expressions, and this in turn may require the introduction of properly designed role-indicators.

4.2. *Analytic vs. synthetic relations*

It was stated above that this distinction seems to be an artificial one; we shall discuss this statement here in more detail. The basis for this distinction seems to lie in the following division of "encoding" (= put in machineable form) "of abstracts" by Melton, 58a:

"Step I. The highly variable phrasing of abstracts written in English or another natural language is reduced to standardized form"; this gives rise to the synthetic relations as expressed by role-indicators.

"Step IL The words or more precisely the terms in the standardized form of abstract are encoded so that important features of meaning are rendered explicit by the code. Also important relationships among terms are similarly recorded in code-form"; this gives rise to analytic relations. Vickery, 59:136, has formulated the distinction as follows: analytic relations are based on the definition of concepts and exist between concepts by virtue of their meaning. Synthetic relations exist between concepts in particular context only (found by experiment or observation).

Now the distinction made by Melton surely has some pragmatic value if applied consciously and cautiously, but it is nevertheless artificial. One abstract may describe a concept in terms of simpler concepts; then this will give rise to a synthetic phrase in the telegraphic abstract. In another abstract the same concept may be referred to by a single term which will appear as such in the telegraphic abstract and later be transformed into an expression of the Semantic Code by means of the analytic relations. Look at the following example given by Melton, 58f:388/89: A search question contains (among other terms) "Beryllium alloy". The search formulation contains the following elements in order to express this concept:

Phrase 1....

Subphrase 2: KEJ. Mat. processed	LALL. 001. Alloy
Subphrase 3: KUJ. component	MATL. 4. BQE Beryllium as metal

Melton goes on to say "subphrases 2 and 3 are the telegraphic expression for the English expression "Beryllium alloy". "Often, however, in literature, alloys are referred to only by a trade-name or a classification, sometimes with no explicit citation of the constituent elements. In such cases the editor of the telegraphic abstract will not always be able to designate these constituent elements and will be obliged to call the alloy by the name given it in the context of the literature.

The code notation for alloys so named, however, will contain the codes of the major alloying elements. For example, the code for Riverside Alloy 32 is:

LALL. METL. 4. BQE. METL. 4. CQU. O32.[1]

Thus another subphrase may be added to the search requirement:

Subphrase 5 KEJ. Material processed LALL. BQE.

Thus the role-indicator KUJ, stating that Beryllium as metal is a component of Riverside Alloy (as stated by the editor of the abstract) and the analytic relation E stating that Riverside Alloy is composed of Beryllium as metal and copper as metal (as stated in the Semantic Code Dictionary) serve exactly the same purpose, and the question remains why two different symbols must be used (the only fact I can see is the complication of the formulation of search questions).

Now to the argument of Vickery, that "analytic relations are based on the definition of concepts". First, I would prefer to state "definition *of terms".* Second, definition is assigning a term to a subject, i.e. to an expression of an IL designed according to our general model. This has terminological aspects which do not concern us here as well as aspects in the idea plane, namely the selection of subjects to which terms have to be assigned. But this in no way whatsoever influences the structure of subjects and the roles to be assigned to their elements. This has been illustrated by our example above; we may give a further one. (Melton, 62, appendix D, page 14): (a) description of subject: "electrons produce a bound which acts upon atoms". If this sentence would appear in an abstract it would be translated into a telegraphic phrase using synthetic role-indicators, (b) Now this subject is called "atomic bound" by definition; the definition as adapted to translation into the Semantic Code would be as follows: *"Atomic bound is* (a connection, specifically a) bound

produced by (particles of physics, specifically) electrons
and *acts upon* (particles of physics, specifically) atoms."

This definition would be translated into the semantic code by means of analytic relations to give the following:

C4NT £ 1004. % PgPH £ 1006. f PFTH £ 1004.

A last remark will show how artificial the distinction between defined and undefined subjects is. We surely would assume that the headings of UDC would belong to the defined subjects. However we are told by Melton, 58b:3_66/67 that in an experiment to make UDC compatible with the Semantic Code, UDC-headings (which consist of several terms) are translated into "miniature encoded abstracts", the relations between UDC-terms being expressed by role-indicators and the UDC-terms themselves being looked up in the Semantic Code Dictionary. There is another aspect in Vickery's argument which shall be discussed in the following section.

4.3. *Paradigmatic vs. syntagmatic relations*

It is stated in SEMA 64/1:43/44 that there are two tasks:

(1) Broadening of the search question by taking into account further related words. This is referred to in Soergel, 65.9, section 1.3 as "extended retrieval mapping". If this is neglected the percentage of relevant documents found remains too low.

[1] Actually Melton gives the following code:

LALL. BQE. CQU. 032 which

is shortened by the application of special rules holding for alloys.

(2) Specification of the search question, the "content words" remaining the same, but the relations being specified.

The problems posed by these tasks are solved by introduction of two types of relations:

(i) "(b) les rapports paradigmatiques ou rapports essentiels et universellement acceptes —par consequence implicites—qui exist entre les Mots" (SEMA, 64/1:7).

(ii) "(c) les rapports syntagmatiques ou rapport contingents et occasioned—par consequence explicites—qui sont suceptibles d'exister entre les mots" (SEMA, 64/1:7). Now this statement, resembling the other aspect of Vickery's argument, mixes up two things which should be thoroughly distinguished, as is done in the following definition:

Synthetic (or syntagmatic) *relations* serve the purpose of combining subjects to give more complex subjects as is described in our general model where synthetic relations in this sense are made up by pairs or triples etc. of role-indicators (as is the case in the telegraphic abstracts of WRU).

It does not matter whether these relations are occasionally or universally accepted, or if the subject is well known and frequently used, or used only once in some publication.

Paradigmatic relations are those which hold between subjects, that is between expressions of the information language according to the rules sketched in our model. (Note that we use interpretation A.2.6(a) of the model, so that the correspondences between terms and subjects, as established in the Semantic Code Dictionary, for example, do not concern as here.) Now, if the equipment is versatile enough, each user may specify his own stipulations and rules for paradigmatic relations, regardless of universal or non-universal acception. He may be aided by a thesaurus, a list of statistical associations or other means, but he is completely free. It may be that equipment restraints impose some pressure on the group of all users to agree on some paradigmatic relations (compare Soergel, 65.9, section 1.0, last section).

Two further remarks may be added:

(1) We may specify a certain type of partial ordering relations A < B by indicating the role which the more general subject B plays in the special subject A. That is, we use the syntagmatic or synthetic role of the simpler subject in the complex subject for the specification of the paradigmatic relation between the complex subject and the more simple subject. Thus we arrive at that aspect of analytic relations of WRU which has perhaps suggested the term "relations". It is quite obvious that it does not matter if one names the syntagmatic role or the accordingly specified paradigmatic relation. But it is certainly confusing if in the same system both ways of naming are introduced so that in fact the same things have double names.

(2) There is a pragmatic differentiation between:

(i) subjects and relations taken up in a thesaurus;

(ii) subjects and relations dealt with in the literature and in search questions and

(ill) subjects and relations which have never been considered (which do not yet have "empirical warrant", such as "influence of the temperature of the Hudson River on the war in Vietnam"). It is well known that some modern documentation systems tend to mix up (i) and (ii) in offering the searcher relations between terms not only from the thesaurus but also from the literature and from search questions already processed.

A.5. *Synthetic relations, roles, and categories*

5.1. If the roles (taken in a wider sense) have no meaning in themselves[1] but only in the frame of some patterns made up of several roles, then we shall call these roles grammatical. It is an additional feature of natural grammar that the same subject may be expressed by such different patterns of roles. If the roles have meanings in themselves, as is usually the case for roles used in ILs, we shall call these roles semantic. Also in this case it is possible of course that there are different ways of expressing the same subject (as for example in the WRU system).

5.2. As outlined in A.2.3. relations (grammatical or semantic) are made up as patterns of several roles. Now it is well known that ILs given in the literature partly use relation-indicators, partly use role-indicators, or both together in an unconscious manner. In order to compare different schemes of roles and/or relations it is therefore necessary to split the relations up into the appropriate roles.

5.3. The roles may also give rise to a categorization (that is a division into disjunct subsets) of primitive or at least fairly simple concepts in the following way: select a subset of roles such that each primitive or simple concept considered may appear in more complex subjects in exactly one role contained in the subset (look for example at the scheme Vi, 59a, given in section B.4.). This then gives a division of concepts according to their possible function rather than by fields of science. Of course there are some roles which are not suited to serve as categories, for example "starting material" and "outcoming material" in the analysis of chemical reactions or "agent" and "target" in the analysis of social relations, because each material or each person may take both of the respective roles.

Such categorized lists of simple concepts are characteristic of faceted classification. A deeper analysis of the concepts contained in a category or facet may reveal that many of these concepts can be reduced to a "role-term" consisting of a "root" and a role-indicator; this occurs definitely in all those cases where the same term occurs in different facets. Thus in faceted classification there may be present a certain degree of prelinking (compare also Soergel, 65.9, section 1.23).

5.4. Finally we should mention the application of roles in the "aid-to-indexing-form" of the American Institute of Physics, or in the "Positions-Referat" (Rothkirch-Trach, 63.3). In both cases one starts with the assumption that some roles should appear in each scientific publication and that a listing of these roles might be a good framework for the abstractor indexer; this would be a specific realization of requirement A. 1.(5). Rothkirch lists the following such roles: Field; Problem to be analyzed; Object which serves for the experiment; Method; Result.

B. COMPARATIVE LIST OF SOME SCHEMES GIVEN IN THE LITERATURE

B .1. *Principles of construction*

1.1. It is of practical importance, for information retrieval, to work out a detailed scheme of role-indicators (see A.3.5.) different sub-sections of which may be used for information retrieval languages in different fields. This task, as indicated in A.2., is an integral part of the broader task of revealing the structure of subjects, that is to design an appropriate classification scheme. The role-indicators may be viewed as a sub-system of such a scheme. They can be broken down into semantic factors or aspects in the same way as "content concepts" can be. There may exist hierarchical relations between role-indicators, but it

[1] It is not stated here that roles by definition have no meaning in themselves; but this case is assumed to be possible.

seems not appropriate to devise too strict a scheme of hierarchy. In short, the role-indicators may be viewed as very general concepts applicable over a broad range of fields of thought.

As a first step to the development of such a universal scheme of role-indicators it seems to be useful to build up a cumulative and comparative list of all schemes designed so far.

In building up such a comparative list, a pragmatic and cautious grouping of role-indicators using the different proposals made by different authors and compromising between these is to be preferred to any artificial stress.[1] A provisional list following these principles is given in B.3., comprising only a small number of sources. Some features evolving from this tentative list are described in B.2. In order to render such a universal scheme really useful, a definition must be attached to each role-indicator; these definitions will in many cases be inductive, that is by a list of examples.

Besides or after the study of role-indicators it is interesting to study the possible patterns of combinations of them. We know already of the synthetic relations as being such combinations which are fixed to a high degree so that one may perhaps say that they lie in the "nature" of the corresponding role-indicators. Other less fixed combinations which are likely to occur are given on "phrase sheets" by Melton, 58a:82f (referred to as possible themes). Furthermore it would be necessary to take into account requirement A.l.II and to state rules of inference from one pattern of role-indicators to another, or from two or more patterns which simultaneously hold to one another (compare SYNTOL 62:32 "regies de developpement"). To give a simple and well known example:

if "a causes b" and "b causes c" then "a causes c"

Requirement A. 1.(5) also gives rise to an analysis of appropriate rules for patterns of roles.

1.2. The building-up of an exhaustive cumulative list will be attempted: a list of all the role-indicators contained in many different sources according to principles described elsewhere (Soergel, 65.5). The full application of these principles would make the use of a computer necessary.

The state of this work is as follows: The role-indicators contained in the sources given in BA have been brought into one list, the different sources being kept, together with some other information. An extract of this list, not giving the different sources, is given as section B.3.

Relations have been split up into their component role-indicators as indicated in A.5. Indirect prelinking has been indicated in the full list.

B.2. *Some features evolving from the comparison of some schemes of role-indicators*

There is a fairly good agreement on some general role-indicators. There are not too many examples of role-indicators contained in only one or two schemes, and in most cases these are special ones.

On the other hand there is a wide variety of proposals for the arrangement and classification of role indicators.[2] I hope that the arrangement I have chosen is in accordance with the principles stated in B.I.I.

[1] As already indicated above, what is a role is defined by enumeration. Since the list is intended to contain all the information from the sources processed, there is no place for selectivity. The hierarchical arrangement is meant only as a tentative and highly pragmatic device for ordering the relations and role-indicators. A final classification which may claim full logical strength will (hopefully) emerge from this approach.

[2] In this regard the situation seems to be not too much different from that in the "content" section ot classification schemes.

Finally it is easy to see that indeed there are not just a few role-indicators which may be •split up into more basic factors. This has been indicated in some few cases. The scheme of "basic factors" has to be developed first in order to render possible a systematic treatment from this problem.

The following might be a first trial Putting aside such very clear things as time and space .and the pairs whole/part, class/element there remain the following basic "types": Thing, entity-type (having a form, Ranganathan's personality) Material, substance-type (having no form, manifesting themselves in thing-entities;

Ranganathan's matter)
Property-type Condition-
type Process-type Agent-
type Target-type

Further aspects would be:

causality; with purpose—without purpose (motivative); real—mental.

The first sequence contains perhaps those roles necessary in each case for the description of phenomena, whereas the further aspects introduce in addition the explanation of phenomena. The distinction real—mental seems to be useful in that the first is prefixed to role-indicators referring to "real" relations between things (as for example action, b strikes c), and the second is prefixed to role-indicators referring to the way in which human thought reveals things and "real" relations between things. It may further be mentioned that many binary relations are used in one possible direction in one scheme and in the other in another scheme (for example part/whole and whole/part) and that it is sometimes difficult to decide from the secondary sources what the direction is.

B.3. *Cumulative and comparative list of the relations and role-indicators contained in the sources given in B.4>*

This table tries to concentrate in an encyclopedic manner the informations contained in each of the sources, that is:

(1) Each piece of information contained in any of the sources must be contained in the cumulative table without any exceptions in such a way that it may be exactly identified.
(2) For each piece of information in the table all the sources must be given.

To achieve this purpose, the following conventions are introduced:

The hierarchical levels are not indicated by indentation but by different typefaces:

COMPARISON
Order relationships
Before
FIRST AMONG
Pa:>yaf

Sometimes the italicizing is neglected for relations or role-indicators having no lower terms.

An identifying mark is assigned to the words or word combinations in each heading; these marks are numbers and sometimes small letters in addition to those in ( ); within

each heading there is a separate count. After each heading those sources are listed which contain a role-indicator grouped under the heading. If a source is not mentioned with some relation or role-indicator, it may be that this source contains in spite of this a more general one or more special relations or role-indicators; this has been indicated in some few, especially important cases. Each source indication is followed (after :) by the notation of the corresponding role-indicator in this source; notation in [ ] means that it has been assigned by the author of this paper (compare B A, introduction). There are three possibilities concerning the terms or descriptions for the relations or role-indicators used by different sources:

(i) The term occurs in the heading (small variations, e.g. singular-plural, are neglected, or, if more important, indicated by :); the source is preceded by the identifying mark of the term.

(ii) The description contains a term occuring in the heading; the source is preceded by the identifying mark of the term, the description (often showing that the role-indicator is more specific or more general, or indicating indirect prelinking) is given in full, the term in question referred to by ….

(iii) There occurs no term of the heading in the source: no identifying mark, description given in full.

Further signs:

If the source in question has an hierarchical arrangement, the position of the role-indicator in the hierarchy is indicated by the following signs:

C > : C has lower terms > CD: there is a higher term for CD

~ : converse, e.g. (d) is the converse of (d)

+: after a notation means: the corresponding role-indicator or relation is broader than the heading concerned (compare (ii) above) "Logical complement"; compare A3.3, footnote 1; "Having the property" would be logical complement of "Property". Indicated only in some few cases.

/…: see also a.c. in: also classified in

The sequence of the sources is as follows:

Schemes of relations
- Ke58
- Pa
- Pe65
- Fa 55
- Ha 57
- Le 59 (incl. categories)
- Mi 51
- Sk61
- WRU relations
- WRU roles
- WRU-Re 58
- WRU-Me 58

Schemes of role-indicators
- WRU see above
- Bo 58
- Co 61
- Ne58
- Wh57
- Wi54

Schemes of categories or of facets

Members of Class. Res. Group
- Vi59a
- Vi60 +
- Ky59
- Ky61
- CRG61

Others

Ar	Pa 62
Ch59a	Ra Re
Ch59b	58 Se56
Co 57	Th51
Ha 64	

BA *Sources of relations, role-indicators and categories*

Most of the schemes have been taken from secondary sources; the following of these are assumed to be easily available:

de Grolier 62
Perreault 65
Vickery 61;

the lists contained therein are only referred to and not duplicated. Lists contained only in other sources or to which major additions had to be made are contained in this section. For each scheme, the different sources where it is reproduced are indicated if known to the author of this report.

The source from which it is cited is indicated by *. If a scheme is not contained in a paper by the resp. author, contained in the list of references, it is denoted by (not bibL).

Notations in [] have been assigned to those schemes which do not have an original notation; this renders reference easier. Of course there arise some difficulties for those schemes which are not duplicated here.

If a scheme is devised for a special field, this field is given.

Ar Aristotle (not bibL): categories
Perreault 65:137
*Vickery 59:212
(slight modifications: accident not mentioned by Vickery; Perreault: position == Vickery: situation; numbers in [] assigned according to Vickery's list)

Be57 Bernier 57: relations among "semantemes" (elements of meaning)
*Vickery 61:27 (a) is on the terminological plane and therefore not considered

Bo58 Booth 58 : 60 "(d) Adverbial [clauses]"

1. time
2. manner
3. place
4. cause
5. comparison

[5a.] degree

6. condition
7. purpose
8. result
9. concession

Some of these terms are also named as "adverbs" on p. 62 and p. 64.
4. cause and 5. purpose are taken together there, and on p. 64 there is further mentioned degree, to which we assign the No, 5a.

Ch59a Cherenin 59/60:403/5: roles (categories)

[M] methods
[O] operations first category
[A] attributes

[Ob] objects
[P] processes basic subjects 2nd category

There are further mentioned
active objects and passive
objects
on p. 405 and it is stated that this is not a rigorous distinction but that there may be a gradual transition between activity and passivity.

Ch59b Cherenin 59/60:417; categories for the division (systematization) of terms
*de Grolier 62:140

Co57 Cordonnier (not bibl.): categories
*de Grolier 62:58

C06I

*Costello 61: Roles (field: chemistry)

Code	Roles
5	Research information, development information on, study of, investigation of
10	Design of, drawing of, design data on, design information on
12	Patent information on (= claim of or disclosure of in a patent)
3	Reactants in a chemical reaction
7	Products of a chemical reaction
5	Reaction medium, atmosphere, solvent, vehicle, dispersion means, environment of, support (chemical, as for a catalyst)
4	Special agent in a reaction or operation of (catalyst, vulcanizing agent, colour stabilizer)
6	Undesired reaction product, reaction by-product, impurity, contaminant, waste, product of a side reaction
1	Using, by means of, by
2	Independent variable studied for its effect on
9	Dependent variable studied for how it is affected
11	Receiving a physical modification (without chemical change)
—	—passively receiving an operation (storage)
—	—passive, possessive, locative (of in, on)
—	—adjectives
—	—proper names, companies, persons etc.
13	Computer program on
14	Market analysis, sale research on, market development information on

(not completely taken)

The same list has been published in Montague 62.4. Essentially the same list with small variations has been published by Vickery 61:49, citing Costello 59 and in -a shortened (and therefore sometimes misleading) form by Perreault *65 :*137, citing Vickery 59.

CRG61 *CRG61:162/63: categories (facets)

1. Real and hypothetical entities. This would include bodies and their parts
2. Properties, subdivided into permanent and temporary

3. Behaviour, processes, effects, subdivided in self-induced and induced by stimuli
4. Intellectual operations such as counting and measuring
5. Physical operations
6. Fields of study and disciplines (when literature can't effectively be classified by previous facets)
7. Mathematics and language
   Compare Ky61 (see below)

Fa Farradane 55: relations
deGrolier 62:94 In the following table
we give informations cited from
column 1: de Grolier 62:94 (perhaps completed by [Coates 60:57/58] and {Vickery 59:78/79})
column 2: *Coates 60:57/58
column 3: * Vickery 59 :78/79
As indicated in all of the named sources Farradane "distinguishes 9 relationships, obtained by correlating two series of characteristics" (de Grolier 62:93). These two series are
1 permanent, 2 temporary, 3 nontime I
distinct, II non-distinct, III concurrent.

Column 1	Column 2	Column 3
11 Causation or functial dependence /influencing or effect relation	/: causes or produces	Agent /; Action raw material /: product
12 Action upon {reaction operator}	/- is the object of action by	agent /- object
13 Distinction from, or substitution for, imitation		
111 Belonging to (genitive) [permanent properties] {appurtenance relation}	/(includes or contains	Object /(constituent object /(property property /(measure
112 'Dimensional' relation [spatial temporal or other]: properties derived [by an object] from environment [temporary properties etc.]	/+ has the environment property of	
113 Equivalence (synonymity)		
III 1 Association (for) also subjective properties III2 Comparison, or relation agent/activity	/; is associated with	
III 3 Co-presence of two otherwise unrelated concepts	*jd* in the presence of	

Ha64 *Harris 64.6 :III-9, figure 1: categories [1] Structural and spatial concepts [2] Time concepts
[3] "Action" (implementation) vocabulary [4] Result [5] Accuracy [6] Efficiency

Ha57 Hayne 57: relation (field: pharmacology) *Vickery59:132
1 action of reference drug on tissue function
2 action of reference drug on metabolism
3 effect of reference drug on action of other drug

Ke58 Kervegant 57/58: relations (proposed for use in UDC) *de Grolier 62:20 Perreault 65:137 Vickery 59:174 (according to Perreault)

Ky61 *Kyle 61 in *CRG61:162/63: facets A Purposes
B Entities
BB Real
BC Hypothetical C Properties and behaviour of A and B
CC Permanent, inherent
CD Temporary, acquired
CF Self-induced
CG Induced by stimuli D Operations on B and C to produce
DD Physical
DF Intellectual
F Products of B, C, D (combined) G Properties and behaviour of F H Activities
HH of F
HJ with, on, through F
JA Systems, study, theories, disciplines, mathematics etc. Relations between B and F and between C and G are obvious; C and G exhibit the feature of indirect prelinking (see A.3.3). This scheme has been developed starting from CRG61 (see above).

Ky59 Kyle 59: categories (field: social sciences) *de Grolier 62:59

Le59 Leroy 59: categories (referred to by numbers in [ ] and relations (referred to by capitals)
Cited from de Grolier 62 :79/80, which cites from Leroy 59 and refers to Braffort 59.9 and Braffort 59.6/61 ADIA. The relations are also reproduced by Vickery 61:50 and in a shortened version by Perreault 65:137.

Mi51 Miller 51: relations
* Vickery 61:29

Ne58 Newman 58: role-indicators (modulants, which inflect roots) (field: technical literature)
*de Grolier 62:125
Vickery 59:25 and 138
Vickery 59/61ICSI: 859
Vickery 61:28

Pa62 *Papier 62:183 (field: technical literature)
[1] fields
[2] objects
[3] processes
[4] properties
[5] substances

Pa Pages (not bibl.): relations (field: social psychology)
*de Grolier 62:72-4
Perreault 65:137
Note that Perreault 65 gives a shortened version and omits the first letter y preceding all codes for relations in Pages' lexicon. There are also some interesting categories contained in the list of main headings of Pages' lexicon given in de Grolier 62:71-72, which, however, have not yet been taken up into our cumulative list.

Pe65 *Perreault 65 :141: roles and relations
(not completely taken)

Ra Ranganathan's (not bibl.) famous categories
personality, matter, energy, space, time are cited on many occasions.

Re58 Rennes, P. 58 (not bibl.): categories hierarchies of content) (field: applied psychology)
*de Grolier 62:68-69
Vickery 61:34
(not separately published according to de Grolier)

Se56 Selye 56/58: roles (field: medicine)
*de Grolier 62:103
Selye 57.9
(not completely taken)

Sk61 Skorokhodko 61: relations (field: design of flying machines heavier than air)
R_1 to be an element of a class
R_2 to have a function
R_3 to be a consequence
R_4
R_5 to create
(R_6 to be subject ?)
R_7 to be an attribute
R_8
R_9 to be a part of

R_{10} to be fixed relative to
R_{tl} to be directed along

R_{13} to be directed perpendicular to
R_{14} to lie in the plane of symmetry

Th51 Thompson 51 (not bibl): categories (field: architecture)
* Vickery 59:7-9
(A) realization (the architectural structure created)
(B) material
(C) processes and problems
(D) geographical place
(E) time
(F) literary form

Wh57 Whaley 57: roles (field: chemistry)
* Vickery 59:132 raw-
(1) material
(2)
(3) Entity casted, surveyed or designed Agent
(4) of adsorption, catalysis, or solution Entity
(5) prepared, fabricated, or analyzed for

Vi Vickery: categories

We divide the various lists given by Vickery into two groups as follows:

Vi59a *Vickery 59 :25, reproduced without any alterations in Vickery 61:34
This is a rather detailed, hierarchical structured list to which we have assigned a notation starting from the notation given by Vickery 60 (see below) and which we therefore reproduce once more.

p Things, substances, entities
PI Naturally occurring
P2 Products
P2:A Tools
P3 Mental constructs
0 Their parts
O1:C Constituents
O2 Organs
P' Systems of things
Q+ El Attributes of things
Q Qualities, properties, including
Ql Structure
Q2 Measures
El Processes, behaviour
R Object of action (patient)
E' Relations between things, interactions
E'l Effects
E'2 Reactions
E3 Operations on things

E31 Experimental
E32 Mental
G: Q' Properties of attributes, relations and operations
G: E3' Operations on attributes, relations and operations
S Place, condition
T Time

Vi60 + Yickery 60 +

Under this heading we have grouped together the following four lists given by Yickery, which differ only slightly from each other.

1. Yi59b:*Vickery59:33
2. Vi59c: Vickery 59: categories or facets (ield: soil science) *Vickery 61:33
3. Vi59d: *Vickery 59-61 ICSI :859 facets
4. Vi60: *Viekery60:30

The following list comprising each term given in any of these four lists together with the different sources, referred to by numbers in () may serve as an example of "encyclopedic cumulation". The notation is taken from Vickery 60:30 (it is not contained in the other sources); if needed notations have been assigned to terms from the other sources. (In accordance with the notation assigned by us to Vi59a).

P substance (3,4), product (3,4), organism (4), thing (1), kinds of soil (2)
O part (all), organ (3,4), structure (2,4) (including parts and layers) (2)
C constituent (all)
Q property (all) and measure (1,2,4)[(1)]
R object of action (4), raw material (4), patient (3)
E action (3,4), operation (all)[(2)], process (all)[(2)], behaviour (4)
A agent (3,4), tool (4)
[Al] apparatus (1)
[A 4- P] substances and equipment used in operations (2)
[E31] laboratory techniques (2)
G general (2,4) property (2?4)? process (4), operation (2,4)
T time (4)
S space (4)

Wi54 Wildhack 54: roles (field: instrumentation)
*Wildhack58:141 Yickery 59:9

(a) condition, physical property, quality, quantity, factor or effect measured (the measurand)
(b) principle on which the instrument operates
(c) object whose condition etc. is to be measured
(d) name of instrument
(e) field of application
(f) stage of instrument development and use
(g) function performed by instrument
(h) bibliographic form of document

[(1)] in (1) [Q2] measure as own category, following property; in (2) property (including measure), [(2)] [El] process and [E3] operation in (1) and (2) separated categories, in (2) followed by in or on soil resp.

WRU58 Western Reserve University (Berry, Kent, Jessica and John Melton, Perry)
Four lists useful for our purposes have been published in *Perry 58 :

1. List of *"analytical relations"* denoted by one capital *Melton 58d: 178/79
de Grolier 62:115/6 (note the sentence directly preceding the list and replace "exists" by "is" wherever *it* appears.) The list is also reproduced by
Vickery 61:27/28 (definitions somewhat rephrased; examples given by Melton 58d included) SEMA64/I:19 (in French; rephrasing of definitions resembles that of Vickery)
Melton 62 :app. d: 13 (without definitions)
Vickery 59:134 Vickery 61:36 Perreault *65:*137

It is planned to integrate the different sources into an "encyclopedic cumulation". Remark: The analytical relations may be interpreted either as role-indicators or as relations (see A4,2).

2. *Role-indicators* (denoted by combinations of 3 capitals) *Melton 58a: 103f., 128-46; complemented by *Rees 58 :176-88
de Grolier 62:112-14 (under the somewhat misleading heading "relationships")

Some comments on this list and corrections of de Grolier's reproduction seem to be appropriate:

"KWV indicated property": the meaning of this role-indicator is better represented by the WRU-formulation "property given". The meaning of this role-indicator is perhaps best represented by the following examples: "Use material of such and such property" or the "effect on material with such and such properties was studied". One may perhaps also say that the "property given" is treated as independent variable.

KUP instead of "specific property": it is better to use the original WRU-formulation "property determined" (measured, produced, .. ♦); to some extent one may say that the "property determined" is treated as a dependent variable,

KXM instead of "negative process": it is better to say "negation (absence) of process".

KIG is "location, geographic" (not organism concerned)

KIB is organization involved (replace here, and in the following role-indicators added by Rees, "organism" by "organization").

3. *Categories* or "non-exclusive headings for a classification of semantic factors to indicate broad relationships"
*Melton 58c: 259-61
de Grolier 62 :118 (without the originally assigned notation) We refer to the comment of de Grolier, but nevertheless think that the scheme has some useful features. We therefore have extracted a scheme of headings; our interpretation of terms follows the general understanding, and not the WRU-usage sometimes difficult to *get* at. So we do enter in our cumulative list under the origin mark WRU-Me58 not the original system of Melton, but a modification thereof which follows.

I. General concepts
- A Ideas (including theory)
- B Relationships (general...)
- C Properties
- D Fields of Endeavour

II. Relationships

III. States

IV. Processes
- B General processes
- A 4- C Physical and material processes
  - D Social (in the widest sense) and industrial processes
  - 2. Constructive processes
  - 3. Destructive processes

V. Substances
- 1. Natural
- 2. Produced by humans (preparations)

VI. Objects
- 1. Natural (includes living beings)
- 2. Products (artifacts, including devices)

4. *Categories to be used in connection with the role-indicators* (p. 104-16, M. 58a). This list is reproduced with some variations in de Grolier 62:114/5.

REFERENCES

	cited in section
BERNIER, C. L. 57 Semantic relations among semantemes. *Am, Doc.* 8(57): 47-50, 211-20	B.4.
BOHNERT, HERBERT G. 63.10/64 English-like systems of mathematical logic for content retrieval. In: *ADI, 26th Annual Meeting,* Chicago 111. 63.10:155-56	A.I.
BOHNERT, HERBERT G. 63.12 Logical-linguistic studies for machine text perusal (Project LOGOS), Semiannual technical status report, covering period May 1» 1963 through Dec. 31, 1963. Yorktown Heights, N.Y.: IBM Th. Watson Res. Ctr. 63.12. (Contract AF49 (638)-1198.) (Short account in Bohnert 63.10/64.)	A.I.
BOOTH, A. D,; BRAND WOOD, L.; CLEAVE, J. R. 58 *Mechanical resolution of linguistic problems.* London/New York: Butterworth/ Academic Press 1958. 306 p.	B.4.
BRAFFORT, P. 59.6/61 ADIA Strategies optimales pour la recherche automatique des informations. In: *Proc, of the ADIA—Conference, Nachr. Dok.*$_t$ Beiheift Nr. 8,	B.4.
BRAEFORT, P.; LEROY, A. 59.9 Des mots-cles aux phrases-eles. In: *Bull BibL France* 4(59) (9):383-91.	B.4.
CHERENIN, V. P.; LAVRENT'EVA A.; ZHIDKOVA, N. V. 59/60 An experimental information language for mechanized searching of scientific-technical literature. In: *Kent* 59/60:389-428	B.4.
COATES, ERIC JAMES 60a *Subject catalogues. Headings and structure.* London: The Library Association 1960. 186 p.	B.4.
CORDONNIER, G. 59/61 Metalanguage pour les traductions d'intercommunicattons entre hommes et son adaptation dans le domaine des machines pour recherches documentaires. In: *Kent* 59/61 :1091-1137 (Ch. 51).	A.I.

cited in section

COSTELLO, J. T. 61.4 BA
Storage and retrieval of chemical research and patent information by links and roles in Du Pont. In: *Am. Doc.* 12(61)2(4): 111-120.

COSTELLO, J. C; WALL, E. 59 BA
Recent improvements in techniques for storing and retrieving information, Wilming ton : Du Pont 1959

CRG61 BA
Classification Research Group Bulletin No. 6. In: J. *Doc,* 17(61)3:156-72.

CROS, R. C; GARDIN, J. C; LEVY, F. 63
Un systems general de documentation automatique, le SYNTOL; Application d la physiologie et aux sciences humaines. Paris: Gauthier-VilJars 1963. 250 p.

DOYLE, LAUREN B. 63.2 A.3.5
The microstatistics of text. Santa Monica, CaL: SDC 63,2.21. 36 p. (= SDC SP-1083,)

FARRADANE, J. E. L. 55 BA
The psychology of classification. In: /. *Doc.* 11(55)4:187-20L

FARRADANE, J. E. L. 52 BA
Further considerations. In: /. *Doc,* 8(1952)2:73-92

FARRADANE, J. E. L. 50 BA
A scientific theory of classification and indexing and its practical applications. In:/. *Doc.* 6(50)2:83-99.

FREEMAN 65 A.l.
Personal communication.

GROLJER, E. DE 62 B.4.
A study of general categories applicable to classification and coding in documentation. Paris: UNESCO 1962. 248 p. (French version I960,)

HARRIS. CLAUDINE 64.6 BA
Dictionary and hierarchy construction in SMART. In: Harvard comp. lab.

HAYNE, R. L.; TURIM, F. 57 BA
Machine retrieval of pharmacological data. In: *Shera* 57:240-49

KENT, A. ed. 59/60/61
Information retrieval and machine translation 1+ II based on the Int. Conf. on Standards on a Common Language for Machine Searching and Translation. New York: Interscience 1960; 1961. (~ Advances in Documentation and Library Science, vol. 3, pt. 1+ II.)

KERVEGANT. D. BA
Subdivision communes de relation; expos6 des motifs mimeo (ct Perreault *65* $_{v}$ note 21)

KERviGANT, D. 57/58(59) BA
Developpement de Fanalyse des relations dans la CDU. In: *Bull. UFOD 26* 4:1-5; also published in *Q. Bull. IAALD* 3(1958): 111-6; and in *Bull, des Bibl de France* (1959) (11) : 502-3; mimeo 57.9.

KOCHEN 65 A.l.
Personal communication.

KYLE, BARBARA 58/59 BA
An examination of some of the problems involved in drafting general classifications and some proposals for their solution. In: *Rev. Doc.* 26(59): 17-21.

LEROY, A.; BRAFFORT, P. 59 B.4
Notice relative a Felaboration d'un codage par phrases cl6s pour la programmations d'un systeme de selection automatique des documents. Saclay: Commissariat a FEnergie Atomique 1959. (— CEA Note 278.)

LEVIEN, ROGER; MARON, M. E. 65.12 A.L
Relational Data File: A tool for mechanized inference execution and data retrieval. Santa Monica, CaL: Rand 65.12. 90 p. (— Rand Memorandum RM-4793-PR.)

MELTON, J. L. 62 A.4.2
The semantic code today. (== AD 263 126; =* AFOSR 1210; shortened version.) In: *Am. Doc.* 13(62)2:176-81.

MELTON, J. 58a AX,AA2,BA,B.1.1
Procedures for preparation of abstracts for encoding. In: *Perry* 58 :69-150.

*

cited in section

MELTON, JESSICA 58b — A.4.2
Encoding of information previously classified. In: *Perry* 58 :361-78.

MELTON, JOHN 58C — B.4.
The semantic code dictionary. In: *Perry* 58 :603-964.

MELTON, JOHN 58d — A.3.5, B.4.
The semantic code. In: *Perry* 58:221-79.

MELTON, JOHN; MELTON, JESSICA S. 58e — A.I.
A method for automatic encoding for languages other than English. In: *Perry* 58:327-40.

MELTON, J. S.; PERRY, J. W. 58f — A.4.2
Introduction to analysis of questions. In: *Perry* 58 :381-456.

MILLER, G. A. 51 — B.4.
Language and communication. New York: 1951.

MONTAGUE, BARBARA A. 62.4 — B.4.
Patent indexing by concept coordination using links and roles [du Pont] *Am, Doc.* 13(62)1:104-11

NEWMAN, S. M. 58 — B.4.
Storage and retrieval of contents of technical literature. Nonchemical information. Washington *D.C.:* US Pat. Off. 1958. 16 p. (= US Pat. Off. Res. Devlpmt. Rept. 12.)

NEWMAN, S. M. 57 — B.4.
Storage and retrieval of contents of technical literature. Nonchemical information. Washington D.C.: US Pat. Off. 1957. 16 p. (= US Pat. Off. Res. Devlpmt, Rept. 4.)

PAPIER, L. S.; CORTELYON, E. 62 — B.4.
Use of a technical word association test in the preparation of a technical thesaurus. In: /. *Doc.* 16(62)4:183-87.

PERREAULT, JEAN M. 65 — B.4.
Categories and relators: A new schema. In: *Rev. Int. Doc.* 32(65)4:136-44.

PERRY, J. W.; KENT, A. 58
Tools for machine literature searching Semantic code dictionary, equipment, procedures.* New York: Interscience 1958. 972 p. (=* Library Science and Docu mentation, vol. 1.)

REES, T. H. 58 — A.3.3, B.4.
Standardized telegraph abstracts from articles in New York Times. In: Perry 58:173-195. (Contract AF 30 (602)-1498.)

ROTHKTRCH-TRACH, K. CHR. GRAF 63.3 — A.5.4
Das Positionsreferat. Brtissel: EURATOM 63.3.18 p. (=* EUR 258.d (Contract No. 028-62-5 CETD).)

SAMMET, JEAN E.; OHLMAN, HERBERT; BOHNERT, H. G. 62.1 — A.O., A.I.
Discussion—The Pro's and Con's of a special IR-language. In: *Comm. ACM* 5(62)1(1): 8-10.

SELYE, H.; MISCHRA, R. K. 57.9 — B.4.
Symbolic shorthand system for medicine and physiology. In: *Fed, Am. Soc. Exp. BioL Proc.* 16(57)3(9): 704-6.

SELYE, H.; NADASKI, M. 56/58 — B.4.
Symbolic shorthand system (SSS)for physiology and medicine. 2nd ed. Montreal: University Institute of Medicine and Experimental Surgery 1958 (1st ed. 1956).

SEMA; EURATOM 64/1 — A.4.3, B.4.
Etudes de documentation automatique, I: Analyse et comparison des langages documentaire. Brussel: EURATOM; Presses Academiques Europeennes 1964. 106 p. (« EUR 409. f, I.) SHERA, J. H.;

KENT, A.; PERRY, J. W. 57
Information systems in documentation. New York: Interscience 1957. (= Advances in documentation and Library Science, vol. 2.)

SKOROKKODKO, E. F. 61 — B.4.
An experiment in the development of a vocabulary of an information language based on step-by-step coding of the meanings of terms. In: *Proc. Conf. IP, MT, Automat. Char. Recogn,*9 N. 8(61): 1-14 (Moscow: Inst. Sc. Info., Ac. Sc. USSR). Translation in JPRS 13056 (21.3.62).

SOERGEL, D. 65.9 — A.2., A.3.1, A.4.3, A.5.3
Mathematical analysis of documentation systems. Freiburg: 1965.9. 61 p. Mimeo.

	cited in section
SOERGEL, D. *65.5* *Outlines of an algorithm for the analysis and the comparison of classification systems.* Freiburg; 1965.5. 30 p., app. Mimeo.	A.3.5, B.1.2
SYNTOL; EURATOM; Association Marc Bloch 62 *Le SYNTOL (Syntagmatic Organization Language); Etude d'un systime general de documentation automatique,* I: Aspects theoriques. Paris: Association Marc Bloch (1962?). 120 p.	B.1.1
UFENSKH 59 The problem of developing a machine language for an information machine. In: *Problemy Kybernetiky* No. 2 (1959) :47f.; translation in: *Problems of Kybernetics* 2(61)356-70.	A.O., A. 1.
VICKERY, B. C. 61 *On retrieval system theory.* London: Butterworths 1961. 159 p. (2nd ed. appeared 1965.)	B.4.
VICKERY, B. C. 60 *Faceted classification. A guide to construction and use of special schemes.* Second impr. ed. London: ASLIB 1960. 70 p.	B,4.
VICKERY, B. C. 59 *Classification and indexing in Science.* 2nd ed. London: Butterworths 1959. 235 p.	A.4.2, A.4.3, B.4.
VICKERY, B. C. 58/59 ICSI Subject analysis for information retrieval. In: *ICSIProc.* 855-865.	B A
WHALEY, F. R. 56.9/57 A deep index for internal technical reports. *Shera 51*:353-83.	B.4.
WILDHACK, W. A.; STERN, J. *5%* The peek-a-boo system—optical coincidence subject cards in information searching. In: *Casey* 58:125-51 (Casey, R. S. *et al.* eds. Punched Cards. New York: Reinhold 1958.)	B.4.
WELDHACK, W. A.; STERN, J. 56.9/57 The peek-a-boo system in the field of instrumentation. In: *Shera* 57:209-26.	B.4.
WILDHACK, W. A.; STERN, J.; SMITH, W. 54.10 Documentation in instrumentation. In: *Am. Doc.* 5(54) (10)223-237).	B.4.

HEILPRIN: The definition of categorization—is it an equivalence class, non-interesting?
FAIRTHORNE: 'Congruence' expresses what Soergel is after.